Description

Method and circuit arrangement for monitoring and, possibly, controlling the transmission capacity of a data transmission path

The invention relates to a method and a circuit arrangement for monitoring and, possibly, for controlling the transmission capacity of a data transmission path which exists between two data devices and is used for transmitting signals, in particular data signals, and on which a primary channel with a relatively low transmission capacity is used, whose magnitude possibly depends on influences which are outside the control of the two data devices, and which can be replaced or added to by means of one or more secondary channels having a relatively high transmission capacity, with test signals being transmitted at specific times or periodically via the primary channel, with the delay time with which these test signals are transmitted being determined, and with a signal which indicates that the primary channel is overloaded being emitted if a specific threshold value is exceeded, in response to which at least one secondary channel can be activated for signal transmission.

The method mentioned above and the circuit arrangement mentioned above take particular account of the fact that the connection of a secondary channel must be initiated by one of the two data devices. In this case, situations can occur in which the relevant connection can be initiated only after the two data devices have been made aware of this by means of appropriate communication, which is handled via the primary channel.

However, in this context, a situation can arise in which the relevant connection - for example for transmitting a large amount of data that has occurred - is urgently required but in which, at the same time, the primary channel is already overloaded since a large amount of data is already being transmitted via it and/or since its capacity has been reduced owing to external influences, which are not under the control of the data terminal. In this case, any communication which is required before the initiation of a connection of the secondary channel between the data devices can not take place, or can take place only with a major delay, so that the urgently necessary connection of the secondary channel cannot take place, or can take place only with a major delay. This then makes the data transmission and the channel utilization very inefficient.

One procedure of the type just mentioned is now known, for example, in conjunction with the transmission of data signals between an ISDN basic access point and a selection node into a network which, as an IP-based network, that is to say a network operating on the basis of an Internet protocol, allows access to the Internet or to an Intranet, and is described in the document "Always On/Dynamic ISDN" by A. Kuzma, Intel Corporation, October 1997, which has been published by the Vendors' ISDN Association on the Internet at http://ww.via-ISDN.org/. This document furthermore describes the measures considered above, which allow data transmission via a data transmission path between a computer or personal computer which is associated with said ISDN basic access point and is also referred to as an AO/DI client (Always On/Dynamic ISDN), and a selection node into the IP-based Internet or Intranet. Such a selection node is also referred to as an AO/DI-PoP (Always On/Dynamic ISDN-Point of Presence).

With regard to the data transmission path mentioned above having the relatively low transmission capacity, it should be noted here that this data transmission path is generally used not only for a connection between just one computer or personal computer and a selection node but, in places, is also at the same time used for a number of such connections, to be precise on a time division basis. With regard to the data signal transmission channels which are available in this way, the expressions logical channels or SVC channels (Switched Virtual Circuits) are also used.

Data signal transmission channels of the type under consideration at the moment are in each case formed in integrated service digital networks (ISDN) within auxiliary channels of a channel arrangement which comprises auxiliary channels and user channels. In the case of an ISDN communications network, which has already been in use for a long time, the auxiliary channel (referred to as the D channel) has a transmission capacity of, for example, 16 kbit/s; the relevant channel arrangement has at least one user channel, but normally two user channels each having a transmission capacity of 64 kbit/s.

Furthermore, a method is known (EP-A-0 905 998) for the transmission of signals, in particular data signals, via an ISDN transmission path between terminals and the Internet. The ISDN transmission path in this case comprises a D channel, which is used as the primary channel, and two B channels, which are used as secondary channels. To allow the primary or D channel to be used appropriately for data signal transmission, this channel must have an adequate transmission capacity and hence bandwidth. In order to find out whether such a transmission capacity and hence bandwidth are available, quality monitoring packets,

which represent test signals, are transmitted either at specific times or periodically between an access server, which connects the ISDN transmission path to the Internet, and a switching center, which is connected to the relevant terminals, via said primary channel. These quality monitoring packets each have a timestamp which makes it possible in said switching center or in the access server to determine the delay time which the respective quality monitoring packet has experienced in the course of being transmitted via the primary channel. If this delay time, which is dependent on the load on the relevant primary channel, exceeds a specific threshold value, the respective terminal which is connected to said switching center is requested to reduce the bandwidth previously used in the primary channel, that is to say to operate with a narrower bandwidth. However, the individual terminal does not in this case have any capability to determine the available transmission capacity, and hence the bandwidth, of the primary channel. However, it is possible to calculate the bandwidth required for a data signal transmission which is to be carried out by a terminal on the basis of the IP packets to be transmitted, in order to use not only the primary channel (D channel) but, possibly, also a secondary channel (B channel) for the relevant data signal transmission. However, neither are any test signals transmitted via the relevant primary channel, nor are their transmission times evaluated, for this process.

A method and an apparatus for determining and taking account of overload situations in a telecommunications switching system are also already known, with said switching center having a central processor for controlling certain operations of the switching center, and having subscriber access line groups and connecting line groups which are connected to the central

processor (US-A-4.511.762). A switching network is connected to the central processor and to said subscriber access line groups and connecting line groups, and this can produce connections between the various subscriber access lines and connecting lines.

If the connecting lines and/or the subscriber access lines are connected to one another, then a message is in each case transmitted to the central processor in the switching system whenever a communication is made via the connecting lines and/or the subscriber access lines and, in addition, an acknowledgement is emitted to the respective line group in response to the relevant message. The finding that the system is overloaded is in this case based on determination of the time after which the central processor receives a response signal to said message. In this case, the respectively measured time is compared with at least one specific threshold value, in order to determine whether the system is overloaded. Even in this case, subscriber circuits which are connected to the individual subscriber access lines generally cannot determine the respectively available transmission capacity from these connection circuits.

The invention is now based on the object of finding a way in which, in a relatively simple manner, it is possible to monitor the transmission capacity of the primary channel of a data transmission path between two data devices such that a requirement for additional transmission capacity for a data signal transmission which is to be carried out can be determined at an early stage, and hence even before data signal transmission, in order to avoid overloading, that is to say so that a secondary channel can be connected in good time.

With regard to a data transmission method of the type mentioned initially, the object described above is achieved according to the invention in that said delay time is determined in that, in response to the emission of the test signals, a response signal is in each case sent back from at least one of the two data devices via the primary channel of said data transmission path to the other data device in response to the reception of the relevant test signals, to said first data device via the primary channel of the data transmission path, which response signal either comprises the respective test signal itself or is a separate signal initiated by it, in that the time interval between the transmission of a test signal by said first data device and the arrival of a response signal which is sent back to it from said other data device is compared with a predetermined threshold value time, which corresponds to a specific current transmission capacity of the primary channel of the data transmission path, forming a comparison result, in response to which a transmission capacity signal is formed, which corresponds to this result and which can be used to activate at least one secondary channel for signal transmission, in that the monitoring of the transmission capacity is carried out deliberately, before transmission of the relevant data, when an amount of data is present which exceeds a defined amount threshold value and is to be transmitted by said first data device to said other data device,
in that the time of the start of deliberate monitoring of the transmission capacity of said data transmission path is used as the point of origin for regular monitoring of the transmission capacity of the relevant data transmission path at time intervals of t,
and in that no further deliberate monitoring of the transmission capacity of said data transmission path is carried out in a situation in which the time period

düakt which has passed since the last monitoring of the transmission capacity is shorter than a defined time period dümin.

The invention results in the advantage that the transmission capacity on said data transmission path can be monitored relatively easily, so that appropriate measures can be taken by the respective data device on the basis of the monitoring result available there and which can indicate, in particular, that the primary channel is overloaded. These measures mean that, in a situation where said transmission capacity signal indicates that said data signal transmission path is overloaded, the relevant data device requests additional transmission capacity in response to said transmission capacity signal. In the case of the channel arrangement at the ISDN basic access point, which was mentioned initially by way of example, the additional transmission capacity can then be provided by requesting at least one user channel or B channel in addition to the primary channel or D channel which has already been used for the data signal transmission, for data signal transmission as a secondary channel, and by also using this for data signal transmission.

Thus, according to the invention, overloading of the data transmission path which is being used can be identified at an early stage, and suitable measures, in particular the connection of a secondary channel, can be initiated immediately. If communication between the terminals via the primary channel is required to do this, then this can thus be carried out before the already overloaded primary channel is also loaded still further by the transmission of data.

It is thus advantageously possible, even before data signal transmission, to determine whether the

transmission capacity which is available for transmission of the relevant data signals on said data transmission path is adequate to avoid overloading. If it is found that the data transmission path will be overloaded when transmitting the data signals which are present, then said first data device can thus request additional transmission capacity before the relevant data signal transmission, thus ensuring problem-free data signal transmission.

Apart from this, the invention ensures that the data transmission path which is normally used for data signal transmission is not unnecessarily loaded by deliberate monitoring immediately following regular monitoring, but can be used virtually immediately for data signal transmission.

Finally, a sensible transition can be made from deliberate monitoring to regular monitoring of the transmission capacity of said data transmission path so as to avoid two monitoring processes occurring unnecessarily at a short time interval after one another.

The transmission of the respective test signal preferably activates a timer which emits an output signal once a defined time interval has elapsed, which output signal, if it occurs before the arrival of said response signal, causes a transmission capacity signal to be emitted which indicates an overload state on said data transmission path. This measure advantageously means that overloading of said data transmission path can even be identified in a situation in which no such response signal arrives at all, or such a response signal arrives only at a time such that the time interval between the transmission of a test signal and the arrival of a response signal sent back in response

to it is greater than the threshold value time mentioned above.

Those signals which are associated with a communication between the two data devices which is necessary for connection of a secondary channel are expediently transmitted at the earliest possible time, that is to say in particular with priority over the transmission of the data which is present. This procedure has the advantage that the time period between identification of the requirement for a secondary channel and the connection of this channel is not unnecessarily lengthened.

In an integrated service digital network (ISDN), in which a switched virtual channel (Switched Virtual Circuit), which in places runs within a D channel, as the primary channel, and at least one B channel is used as the secondary channel, message signals based on a bandwidth allocation protocol are used for allocating the bandwidth and transmission capacity to be used before setting up a B channel, and are transmitted with priority over the other data. Efficient data signal transmission can thus be carried out in an advantageous manner between the two said data devices in an integrated service digital communications network. Signals from an existing bandwidth allocation protocol can thus easily be used here.

The message signals EchoRequest and EchoReply of an Internet link control protocol are expediently used as the test signal and response signal, respectively. This results in the advantage that it is possible to use signals in accordance with a transmission protocol which is used in any case.

In order to carry out the method according to the invention, it is possible to use a circuit arrangement which is characterized in that at least one of two data devices which are connected to one another via a data transmission path has an associated monitoring device, which allows a time comparison to be carried out between a measurement time interval from the emission of a test signal from the relevant data device to the other data device until the arrival of a response signal from this other data device with a predetermined threshold value time, and in that the relevant monitoring device can emit a transmission capacity signal which corresponds to the respective time comparison result, in particular a report signal which indicates an overload state on the data transmission path, if said measurement time interval exceeds the relevant threshold value time.

This circuit arrangement is distinguished by the advantage of particularly low circuit complexity.

A timer is expediently connected to said monitoring device, can be activated by said test signal, and, emits an output signal to the relevant monitoring device once its operating time, which corresponds to an overload state of said data transmission path has elapsed, which monitoring device uses this output signal, if the response signal from said other data device has not yet arrived, to emit a report signal which indicates the overload state of the data transmission part. This ensures, with particularly low circuit complexity, that overloading of said data transmission path can be identified even in a situation in which said response signal does not arrive at all, or arrives only at a time such that the time interval between the transmission of a test signal and the arrival of a response signal sent back in response to

it is greater than the threshold value time mentioned above.

Before explaining the invention further with reference to an example, it should first of all also be noted that the method according to the invention has a specific application in an integrated service digital network, a so-called ISDN, as is specified by the ITU-T Series 1 Recommendations. In this case, a so-called D channel with a maximum of 16 kbit/s and two B channels each having 64 bit/s are available